

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



RESULT 2  
US-09-105-678A-47  
Sequence 47, Application US/09105678A  
; Sequence 47, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Sueyaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REFERENCE/DOCKET NUMBER: 27,026  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-48  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-48  
; QUERY MATCH  
; QUERY Match 96.3%; Score 105; DB 3; Length 21;  
; Best Local Similarity 90.0%; Pred. No. 1.5e-11; 2; Indels 0; Gaps 0;  
; Matches 18; Conservative 0; Mismatches 2;  
; INQUIRY  
; Qy 1 TPDINPAWXXRGIRPVGRF 20  
; Db 1 TPDINPAWASRGIRPVGRF 20  
; RESULT 4  
; US-09-105-678A-9  
; Sequence 9, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Sueyaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.



ADDRESSE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:

TOPOLGY: linear  
MOLECULE TYPE: peptide  
US-09-105,678A-34

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:

TOPOLGY: linear  
MOLECULE TYPE: peptide  
US-09-105,678A-34

Query Match 96.3%; Score 105; DB 3; Length 33;  
Best Local Similarity 90.0%; Pred. No. 2.6e-11;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy I TPDINPAWYXXRGIRPVGRF 20  
Db 12 TPDINPAWYASRGIRPVGRF 31

Query Match 95.4%; Score 104; DB 3; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2.2e-11;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy I TPDINPAWYXXRGIRPVGRF 20  
Db 1 TPDINPAWYASRGIRPVGRF 20

Query Match 95.4%; Score 104; DB 3; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2.2e-11;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy I TPDINPAWYXXRGIRPVGRF 20  
Db 12 TPDINPAWYASRGIRPVGRF 31

RESULT 8  
US-09-105,678A-34  
; Sequence 34, Application US/09105678A  
; Patent No. 6,03882  
; GENERAL INFORMATION:

APPLICANT: Suehaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko

APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:  
ADDRESSE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:

TOPOLGY: linear  
MOLECULE TYPE: peptide  
US-09-105,678A-40

Query Match 95.4%; Score 104; DB 3; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2.2e-11;

APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

RESULT 10  
 US-09-105-678A-35  
 Sequence 35, Application US/09105678A  
 Patent No. 6103882  
 GENERAL INFORMATION:  
 APPLICANT: Suenaga, Masato  
 APPLICANT: Moriya, Takeo  
 APPLICANT: Tanaka, Yoko  
 APPLICANT: Nishimura, Osamu  
 TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 NUMBER OF SEQUENCES: 52

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30

APPLICATION NUMBER: US/09/105,678A  
 FILING DATE: 26-JUN-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 172118/1997  
 FILING DATE: 27-JUN-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gorlin, David G.  
 REGISTRATION NUMBER: 27,026  
 REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 35:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 21 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide

US-09-105-678A-35

Query Match 95.4%; Score 104; DB 3; Length 21;  
 Best Local Similarity 90.0%; Pred. No. 2.3e-11; 0; Mismatches 2; Indels 0;  
 Matches 18; Conservative 0; MisMatches 2; Indels 0; Gaps 0;

RESULT 11  
 US-09-105-678A-35  
 Sequence 36, Application US/09105678A  
 Patent No. 6103882  
 GENERAL INFORMATION:  
 APPLICANT: Suenaga, Masato  
 APPLICANT: Moriya, Takeo  
 APPLICANT: Tanaka, Yoko  
 APPLICANT: Nishimura, Osamu  
 TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A  
 FILING DATE: 26-JUN-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 172118/1997  
 FILING DATE: 27-JUN-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Gorlin, David G.  
 REGISTRATION NUMBER: 27,026  
 REFERENCE/DOCKET NUMBER: 48466-342  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-523-3400  
 TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 41:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 21 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide

US-09-105-678A-41

Query Match 95.4%; Score 104; DB 3; Length 21;  
 Best Local Similarity 90.0%; Pred. No. 2.3e-11; 0; Mismatches 2; Indels 0;  
 Matches 18; Conservative 0; MisMatches 2; Indels 0; Gaps 0;

RESULT 12  
 US-09-105-678A-36  
 Sequence 36, Application US/09105678A  
 Patent No. 6103882  
 GENERAL INFORMATION:  
 APPLICANT: Suenaga, Masato  
 APPLICANT: Moriya, Takeo  
 APPLICANT: Tanaka, Yoko  
 APPLICANT: Nishimura, Osamu  
 TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/105,678A  
 FILING DATE: 26-JUN-1998  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 172118/1997  
 FILING DATE: 27-JUN-1997

## ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:

LENGTH: 22 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLogy: linear

MOLECULE TYPE: peptide

US-09-105-678A-36

Query Match

Best Local Similarity 90.0%; Pred. No. 2.4e-11; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Query Match

95.4%; Score 104; DB 3; Length 22;

RESULT 13

US-09-105-678A-42

; Sequence 42, Application US/09105678A

; Patent No. 6103882

GENERAL INFORMATION:

; APPLICANT: Sueuaga, Masato

; APPLICANT: Moriya, Takeo

; APPLICANT: Tanaka, Yoko

; APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS &amp; CUSHMAN, LLP

; STREET: 130 Water Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/105,678A

; FILING DATE: 26-JUN-1998

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 172118/1997

; FILING DATE: 27-JUN-1997

; ATTORNEY/AGENT INFORMATION:

; NAME: Conlin, David G.

; REGISTRATION NUMBER: 27,026

; REFERENCE/DOCKET NUMBER: 48466-342

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-523-3400

; TELEFAX: 617-523-6440

; INFORMATION FOR SEQ ID NO: 42:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 31 amino acids

; TYPE: amino-acid

; STRANDEDNESS:

; TOPOLogy: linear

; MOLECULE TYPE: peptide

; US-09-105-678A-7

RESULT 15

US-09-105-678A-8

; Sequence 8, Application US/09105678A

; Patent No. 6103882

GENERAL INFORMATION:

; APPLICANT: Sueuaga, Masato

; APPLICANT: Moriya, Takeo

; APPLICANT: Tanaka, Yoko

; APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

Best Local Similarity 90.0%; Pred. No. 2.4e-11; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2;

Indels 0; Gaps 0;

Query Match

95.4%; Score 104; DB 3; Length 22;

NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 1771187/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGN: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
us-09-105-678A-B

Query Match 95.4%; Score 104; DB 3; Length 31;  
Best Local Similarity 90.0%; Bred. No. 3.5e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPDINPAWYXRGGRPVGRF 20  
Db 12 TPDINPAWYXRGGRPVGRF 31

Search completed: April 17, 2001, 15:39:49  
Job time: 318 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
 Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 17, 2001, 15:38:42 ; Search time 116.94 Seconds  
 (without alignments)  
 10.754 Million cell updates/sec

DM Protein - Protein search, using sw model

Title: US-09-446-543a-73  
 Perfect score: 109  
 Sequence: 1 TPDINPAWXXRGIRPVGRFXX 22

Scoring table: BLOSUM62  
 Gapov 10.0 , Gapext 0.5

Searched: 390729 seqs, 5716235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0  
 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : A\_Genesed\_0401:\*

1: /SIDS1/gcdata/geneseq/geneseq/genesedp/AA1980.DAT:\*

2: /SIDS1/gcdata/geneseq/genesedp/AA1981.DAT:\*

3: /SIDS1/gcdata/geneseq/genesedp/AA1982.DAT:\*

4: /SIDS1/gcdata/geneseq/genesedp/AA1983.DAT:\*

5: /SIDS1/gcdata/geneseq/genesedp/AA1984.DAT:\*

6: /SIDS1/gcdata/geneseq/genesedp/AA1985.DAT:\*

7: /SIDS1/gcdata/geneseq/genesedp/AA1986.DAT:\*

8: /SIDS1/gcdata/geneseq/genesedp/AA1987.DAT:\*

9: /SIDS1/gcdata/geneseq/genesedp/AA1988.DAT:\*

10: /SIDS1/gcdata/geneseq/genesedp/AA1990.DAT:\*

11: /SIDS1/gcdata/geneseq/genesedp/AA1991.DAT:\*

12: /SIDS1/gcdata/geneseq/genesedp/AA1992.DAT:\*

13: /SIDS1/gcdata/geneseq/genesedp/AA1993.DAT:\*

14: /SIDS1/gcdata/geneseq/genesedp/AA1994.DAT:\*

15: /SIDS1/gcdata/geneseq/genesedp/AA1995.DAT:\*

16: /SIDS1/gcdata/geneseq/genesedp/AA1996.DAT:\*

17: /SIDS1/gcdata/geneseq/genesedp/AA1997.DAT:\*

18: /SIDS1/gcdata/geneseq/genesedp/AA1998.DAT:\*

19: /SIDS1/gcdata/geneseq/genesedp/AA1999.DAT:\*

20: /SIDS1/gcdata/geneseq/genesedp/AA2000.DAT:\*

21: /SIDS1/gcdata/geneseq/genesedp/AA2001.DAT:\*

22: /SIDS1/gcdata/geneseq/genesedp/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

12 105 96.3 31 21 B10352 Human oxytocin sec  
 13 105 96.3 31 21 Y49291 19P2 ligand peptid  
 14 105 96.3 32 18 W31392 Human type G prote  
 15 105 96.3 32 21 B10363 Human oxytocin sec  
 16 105 96.3 33 18 W31393 Human type G prote  
 17 105 96.3 33 21 B10364 Human oxytocin sec  
 18 105 96.3 37 18 W31390 Human oxytocin sec  
 19 105 96.3 37 20 W97226 Human type G prote  
 20 105 96.3 37 21 B10361 Human oxytocin sec  
 21 104 95.4 20 18 W31387 Bovine oxytocin  
 22 104 95.4 20 18 W31374 Bovine G protein-c  
 23 104 95.4 20 20 W97232 Bovine pituitary-d  
 24 104 95.4 20 20 W97234 Rat type G prote  
 25 104 95.4 20 20 W95191 Bovine pituitary-d  
 26 104 95.4 20 20 W95175 Murine oxytocin  
 27 104 95.4 20 21 B10350 Bovine oxytocin se  
 28 104 95.4 20 21 B10358 Rat oxytocin secre  
 29 104 95.4 20 21 Y49301 19P2 ligand peptid  
 30 104 95.4 20 21 Y49302 Rat oxytocin secre  
 31 104 95.4 21 18 W31388 Bovine G protein-c  
 32 104 95.4 21 18 W31376 Rat type G protein  
 33 104 95.4 21 20 W95193 Bovine pituitary-d  
 34 104 95.4 21 21 B10351 Bovine oxytocin se  
 35 104 95.4 21 21 B10359 Rat oxytocin secre  
 36 104 95.4 22 18 W31389 Bovine G protein-c  
 37 104 95.4 22 18 W31376 Rat type G protein  
 38 104 95.4 22 20 W95193 Bovine pituitary-d  
 39 104 95.4 22 21 B10352 Bovine oxytocin se  
 40 104 95.4 22 21 B10360 Rat oxytocin secre  
 41 104 95.4 31 18 W31384 Rat type G protein  
 42 104 95.4 31 18 W31371 Bovine G protein-c  
 43 104 95.4 31 20 W97233 Rat type ligand po  
 44 104 95.4 31 20 W97233 Bovine pituitary-d  
 45 104 95.4 31 21 Y87504 Rat prolactin rele

**ALIGNMENTS**

RESULT 1  
 W31394 ID W31394 standard; Peptide; 20 AA.  
 XX  
 W31394: AC  
 XX  
 DT 06-APR-1998 (first entry)  
 XX  
 DE Human type G protein-coupled receptor ligand fragment 4.  
 XX  
 KW G protein-coupled receptor; ligand binding; pharmaceutical;  
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0972436-A2.  
 XX  
 PD 10-JUL-1997.  
 XX  
 PF 26-DEC-1996; 96WO-JP03821.  
 XX  
 PR 18-SEP-1996; 96JP-0246573.  
 PR 28-DEC-1995; 95JP-0343311.  
 PR 15-MAR-1995; 96JP-0059419.  
 PR 12-AUG-1996; 96JP-0211805.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawamura Y, Kiada C;  
 XX  
 DR WPI; 1997-363672/33.  
 DR N-PSDB; V02431.

PT Ligand peptide for G protein-coupled receptor - acts by modulating function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS  
 XX  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
 CC sequence represented in W31390 and is used in an assay to monitor ligand  
 CC binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a prophylactic  
 CC or therapeutic agent for dementia, depression, hypokinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,  
 CC hypercholesterolaemia, hyperglycaidaemia, Hyperprolactinaemia, diabetes,  
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,  
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 CC osteoporosis and/or oligogalactia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the  
 CC ligand affecting activation of the G protein-coupled receptor protein.  
 XX

80 Sequence 20 AA;

Query Match 96.3%; Score 105; DB 18; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 1.3e-11;  
 Matches 18; Conservative 0; Mismatches 2; Index 0; Gaps 0  
 Qy . 1 TPDINPANWXXRGTRPVGSRP 20  
 Db 1 tpdinpanwqasrdipvgif 20

RESULT 2 .

XX ID W97236 standard; peptide; 20 AA.  
 XX AC W97236;  
 XX DT 06-MAY-1999 (first entry)

XX DE Human type ligand polypeptide fragment.

XX H H Rat type ligand; modulation; proactin secretion; G protein-coupled receptor; GPCR; hypoovarianism; gonocyst carcinogenesis; menopausal syndrome; embryo; brain tumour; hypometabolism; lactation; pituitary adenomatisis; infertility; impotence; amenorrhea; galactorrhea; proactinoma; acromegaly; Chari-Frommei syndrome; Argonz-del Castillo syndrome; Forbes-Albright syndrome; lymphoma; Sheehan's syndrome; dyszoospermia; contraceptive; placental function; choriocarcinoma; hydatid mole; abortion; abortifacient; abortifacient; abortifacient; abortifacient; abortion mole; abortion; unthrifly fetus; abnormal saccharometabolism; abnormal lipidmetabolism; oxytocia.  
 XX OS Homo sapiens.  
 XX PN W09858662-A1.  
 XX PD 30-DEC-1998.  
 XX PP 22-JUN-1998; 98W0-JP02765.  
 PR XX 23-JUN-1997; 97JP-0165437.  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 PI Fujii, R., Hinuma, S., Kawamata, Y., Matsumoto, H.;  
 XX WPI; 1999-105614/09.  
 DR

PT Use of G Protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 PS Claim 3: Page 166: 241bp; English.  
 SQ The present sequence represents a human type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand Polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypoparathyroidism, gonad cysticacogenesis, menopausal  
 CC syndrome, euthyroid or hypothyroidism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, eunuchopathy, autoimmune disease,  
 CC proctomatoma, infertility, impotence, amenorrhea, galactorrhea,  
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyzoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC choriocarcinoma, hydatid mole, interruption more, abortion, unthrifty fetus,  
 CC abnormal saccharonametabolism, abnormal lipidmetabolism or oxytocia.  
 XX Sequence 20 AA:  
 Query Match 96.3%; Score 105; DB 20; Length 20;  
 Best Local Similarity 90.0%; pred. No. 1.3e-11;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TPDINPAWYXXRGIRPVGHR 20  
 DB 1 tpoinpawasrgirpvgrf 20  
 AC B10365;  
 AC B10365;  
 RESULT 3  
 XX DT 24-NOV-2000 (first entry)  
 XX DE Human oxytocin secretion promoting peptide SEQ ID NO: 35.  
 XX KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 XX treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 XX caesarean section; artificial fertilization; galactostasis; goat; pig;  
 XX veterinary medicine; milk production.  
 XX OS Homo sapiens.  
 XX WO200038704-A1.  
 XX PD 06-JUL-2000.  
 XX PF 22-DEC-1999; 99W0-JP07199.  
 XX PR 25-DEC-1998; 98JP-0369585.  
 XX PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX PI Matsumoto H, Kitada C, Hinuma S;  
 XX DR WPI; 2000-452298/39.  
 XX PT Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -  
 XX

PS Disclosure; Page 63; 72pp; Japanese.

XX

CC This invention describes a novel oxytocin secretion-regulating agent which contains a ligand peptide or its salt for the G protein-coupled receptor protein. It is useful in the form of drugs for ameliorating, preventing and treating diseases relating to oxytocin secretion e.g., weak pains and atonic bleeding, before and after expulsion of placenta, uterine recovery failure, caesarean section, stoppage of artificial fertilization or galactostasis and is also applicable in veterinary medicine for promoting milk production in cow, goat and pig. This sequence represents a human peptide which acts as an oxytocin secretion promoter.

CC

XX

SQ Sequence 20 AA;

Query Match	Score	DB	Length
Best Local Similarity	96.3%	21	20
Matches	18;	Conservative	0;
ID	Y49294	Pred. No.	1.3e-11;
AC	Y49294;	Mismatches	2;
XX		Indels	0;
XX		Gaps	0;

QY 1 TPDINPAWYXRGIRPVGRF 20

Db 1 tpdinpawyasrgirpvgrf 20

RESULT 4

Y49294 standard; peptide; 20 AA.

XX

AC Y49294;

XX

XX 22-FEB-2000 (first entry)

DE 19P2 ligand peptide fragment.

XX

KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion; pituitary; regulatory mechanism; central nervous system; pancreatic.

XX

OS Homo sapiens.

XX

KEY Location/Qualifiers

PH Modified-site 20

PT /note= "C-terminal amide"

XX

PN WO960112-A1.

XX

PD 25-NOV-1999.

XX

PT 20-MAY-1999; 99WO-JP02650.

XX

PR 21-MAY-1998; 98JP-0140293.

XX

PA (TAKE ) TAKEDA CHEM IND LTD.

XX

PT Matsuoto H, Kitaada C, Hinuma S;

PI Fujii R, Fukumori S, Habbata Y, Hinuma S, Hosoya M;

PT Kawamata Y, Kitada C;

XX

PR WPI: 1997-353672/33.

XX

DR N-PSDB; V02432.

XX

PT Ligand peptide for G protein-coupled receptor - acts by modulating function in the central nervous system, pancreas and pituitary gland

XX

Claim 2; Page 186; 258pp; English.

XX

CC This sequence represents a peptide fragment from a novel human type ligand polypeptide corresponding to amino acid residues 34 to 54 of the sequence represented in W31390 and is used in an assay to monitor ligand binding to the G protein-coupled receptor protein. Pharmaceutical compositions containing this ligand may be used as a pituitary function modulator, a central nervous system modulator or a pancreatic function modulator. This ligand could have specific applications as a prophylactic or therapeutic agent for dementia, depression, hyperkinetic syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone secretery disease, hyper- and polyphagia, hyperlipidaemia, hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis, acute myocardial infarction, infertility, spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis and/or oligogalactia. Assays can also be developed to screen compounds which are capable of altering the binding activity of the ligand affecting activation of the G protein-coupled receptor protein.

XX

Sequence 21 AA;

Query Match	Score	DB	Length
Best Local Similarity	96.3%	21	20
Matches	18;	Conservative	0;
ID	W31395	Pred. No.	1.3e-11;
AC	W31395;	Mismatches	2;
XX		Indels	0;
XX		Gaps	0;

QY 1 TPDINPAWYXRGIRPVGRF 20

Db 1 tpdinpawyasrgirpvgrf 20

RESULT 5

W31395

ID W31395 standard; Peptide; 21 AA.

XX

AC W31395;

XX

DT 06-APR-1998 (first entry)

DE Human type G protein-coupled receptor ligand fragment 5.

XX

KW G protein-coupled receptor; ligand binding; pharmaceutical; modulator; pituitary; central nervous system; pancreas; prophylactic; therapeutic agent.

XX

PT 10-JUL-1997.

XX

PR 26-DEC-1996; 96WO-JP03821.

XX

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059119.

PR 12-AUG-1996; 96JP-0211805.

XX

PA (TAKE ) TAKEDA CHEM IND LTD.

XX

PT Fujii R, Fukumori S, Habbata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

XX

PR WPI: 1997-353672/33.

XX

DR N-PSDB; V02432.

XX

PT Ligand peptide for G protein-coupled receptor - acts by modulating function in the central nervous system, pancreas and pituitary gland

XX

Claim 2; Page 186; 258pp; English.

XX

CC This sequence represents a peptide fragment from a novel human type ligand polypeptide corresponding to amino acid residues 34 to 54 of the sequence represented in W31390 and is used in an assay to monitor ligand binding to the G protein-coupled receptor protein. Pharmaceutical compositions containing this ligand may be used as a pituitary function modulator, a central nervous system modulator or a pancreatic function modulator. This ligand could have specific applications as a prophylactic or therapeutic agent for dementia, depression, hyperkinetic syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone secretery disease, hyper- and polyphagia, hyperlipidaemia, hypercholesterolaemia, hyperglycaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis, acute myocardial infarction, infertility, spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis and/or oligogalactia. Assays can also be developed to screen compounds which are capable of altering the binding activity of the ligand affecting activation of the G protein-coupled receptor protein.

XX

Sequence 21 AA;





KW irrigation mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9858962-A1.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PR 22-JUN-1998; 98WO-JP02765.  
 XX  
 PR 23-JUN-1997; 97JP-0165437.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PT Fujii R, Hlnuma S, Kawamata Y, Matsumoto H;  
 XX  
 DR WPI: 1999-105614/09.

XX  
 PT Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 PS Claim 3; Page 159; 241PP; English.

XX  
 CC The present sequence represents a human type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypoovaraism, gonocyst carcinogenesis, menopausal  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatis, brain tumour, ameniopathy, autoimmune disease,  
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyzoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
 XX  
 SQ Sequence 31 AA;

Query Match 96.3%; Score 105; DB 20; Length 31;  
 Best Local Similarity 90.0%; Pred. No. 2.1e-11; Mismatches 0;  
 Matches 18; Conservative 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYXRGIRPVGRF 20  
 Db 12 tpdinpawysrgirpvgrf 31

RESULT 11  
 W87615  
 ID W87615 standard; Peptide: 31 AA.  
 AC W87615;  
 XX  
 DT 29-MAR-1999 (first entry)  
 XX  
 DE Human 19P2 ligand.  
 XX  
 KW 19P2 ligand; G protein coupled receptor; pituitary;  
 KW prolactin releasing peptide; human; dementia; breast cancer;  
 KW therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN EP87417-A2.

Query Match 96.3%; Score 105; DB 20; Length 31;  
 Best Local Similarity 90.0%; Pred. No. 2.1e-11; Mismatches 0;  
 Matches 18; Conservative 0; Indels 2; Gaps 0;

Qy 1 TPDINPAWYXRGIRPVGRF 20  
 Db 12 tpdinpawysrgirpvgrf 31

RESULT 12  
 B10362  
 ID 810362 standard; peptide: 31 AA.  
 XX  
 AC B10362;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Human oxytocin secretion promoting Peptide SEQ ID NO: 32.  
 XX  
 KW Human oxytocin secretion promoting Peptide Protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.

XX  
 PD 30-DEC-1998.  
 XX  
 PR 25-JUN-1998; 98EP-0111725.  
 XX  
 PR 27-JUN-1997; 97JP-0172118.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PT Moriya T, Nishimura O, Suenaga M, Tanaka Y;  
 XX  
 DR WPI; 1999-047884/05.  
 XX  
 PT producing a 19P2 pituitary G protein receptor ligand - by cleavage  
 PT of a fusion protein, useful for preventing and treating dementia,  
 PT breast cancer, renal failure and autoimmune disease  
 XX  
 PS Claim 5; Page 35; 56pp; English.

XX  
 CC This is the amino acid sequence of the human pituitary G  
 CC protein-coupled receptor ligand 19P2L. A method suitable for  
 CC commercial high-level production of 19P2L comprises expressing  
 CC the ligand in host cells as a recombinant fusion protein e.g. with  
 CC human basic fibroblast growth factor (see WO979697) that has  
 CC been modified to include an N-terminal cysteine residue. The  
 CC ligand is released from the fusion by cyanylation followed by  
 CC ammonolysis. 19P2L has prolactin secretion-stimulating and (at  
 CC high doses) prolactin secretion-inhibiting properties. It can be  
 CC used in the treatment and prevention of various diseases, including:  
 CC senile dementia, cerebrovascular dementia, and dementia associated  
 CC with: genetic disorders (e.g. Alzheimer's disease, Parkinson's  
 CC disease, Pick's disease, Huntington's disease), infectious diseases  
 CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or  
 CC toxicosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,  
 CC intoxication by drugs, metal and organic compounds), tumourigenic  
 CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic  
 CC subarachnoidal hemorrhage, and other types of dementia, depression,  
 CC hyperactive child syndrome (microencephalopathy) and disturbance of  
 CC consciousness. It is also useful for prevention and treatment of  
 CC diseases associated with prolactin hypo and hypersecretion  
 CC respectively, including: hyperprolactinemia, pituitary adenoma,  
 CC breast cancer, infertility, impotence and autoimmune disease  
 CC (hyposecretion disorders), and seminal vesicle hypoplasia,  
 CC osteoporosis, menopausal syndrome and renal failure (hyposecretion  
 CC disorders). The 19P2 polypeptide/amide is also useful as a test  
 reagent for study of the prolactin secretory function or as a  
 CC lactogogue in mammalian farm animals.  
 XX  
 SQ Sequence 31 AA;

Query Match 96.3%; Score 105; DB 20; Length 31;  
 Best Local Similarity 90.0%; Pred. No. 2.1e-11; Mismatches 0;  
 Matches 18; Conservative 0; Indels 2; Gaps 0;

Qy 1 TPDINPAWYXRGIRPVGRF 20  
 Db 12 tpdinpawysrgirpvgrf 31

RESULT 12  
 B10362  
 ID 810362 standard; peptide: 31 AA.  
 XX  
 AC B10362;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Human oxytocin secretion promoting Peptide SEQ ID NO: 32.  
 XX  
 KW Human oxytocin secretion promoting Peptide Protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.

XX  
OS Homo sapiens.  
XX  
PN WO200038794-A1.  
XX  
PD 06-JUL-2000.  
XX  
PF 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98JP-0360585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
DR WPI; 2000-039381/03.  
XX  
PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
studying diseases related to ligand abnormality.  
XX  
PS Disclosure; Page 26; 73pp; Japanese.  
XX  
CC This invention describes a novel oxytocin secretion-regulating agent  
which contains a ligand peptide or its salt for the G protein-coupled  
receptor protein. It is useful in the form of drugs for ameliorating,  
preventing and treating diseases relating to oxytocin secretion e.g.,  
weak pains and atonic bleeding, before and after expulsion of placenta,  
uterine recovery failure, caesarean section, stoppage of artificial  
fertilization or galactostasis and is also applicable in veterinary  
medicine for promoting milk production in cow, goat and pig. This  
sequence represents a human peptide which acts as an oxytocin secretion  
promoter.  
XX  
SO Sequence 31 AA;  
Query Match 96.3%; Score 105; DB 21; Length 31;  
Best Local Similarity 90.0%; Pred. No. 2 1e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 TPPDINPAWYXRGIRPVGRF 20  
Db 12 tpdinpawyasrgirpvgrf 31  
RESULT 14  
ID W31392  
ID W31392 standard; Peptide; 32 AA.  
XX  
AC W31392;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 2.  
XX  
KW G Protein-coupled receptor; ligand binding; pharmaceutical;  
modulator; pituitary; central nervous system; pancreas; prophylactic;  
therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN W0972436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-033371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211895.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
XX  
DR WPI; 1997-363672/33.  
XX  
N-PSDB; V02429.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 185; 258pp; English.

XX  
OS Homo sapiens.  
XX  
PN WO990112-A1.  
XX  
PD 25-NOV-1999.  
XX  
PR 20-MAY-1999; 99WO-JP02650.  
XX  
PR 21-MAY-1998; 98JP-0140293.

XX  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 34 of the  
 CC sequence represented in WO1390 and is used in an assay to monitor ligand  
 CC binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a prophylactic  
 CC or therapeutic agent for dementia, depression, hyperkinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia,  
 CC hypercholesterolaemia, hyperthyroidism, hypopituitarism, diabetes,  
 CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis, asthma,  
 CC rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy,  
 CC amyotrophic lateral sclerosis, acute myocardial infarction, infertility,  
 CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
 CC osteoporosis and/or oligoplasia. Assays can also be developed to screen  
 CC compounds which are capable of altering the binding activity of the G protein-coupled receptor protein.  
 XX  
 SQ Sequence 32 AA:

Query Match 96.3%; Score 105; DB 18; Length 32;  
 Best Local Similarity 90.0%; Pred. No. 2.2e-11; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TPDINPAWYXKGIRPGRF 20  
 Db 12 tpdinpawyasrgirpgvgrf 31

Search completed: April 17, 2001, 15:38:42  
 Job time: 349 sec

CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.  
 XX  
 SQ Sequence 32 AA:

Query Match 96.3%; Score 105; DB 21; Length 32;  
 Best Local Similarity 90.0%; Pred. No. 2.2e-11; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TPDINPAWYXKGIRPGRF 20  
 Db 12 tpdinpawyasrgirpgvgrf 31

RESULT 15  
 B10353 ID B10363 standard; peptide; 32 AA.  
 XX  
 AC B10363;  
 XX  
 DT 24-NOV-2000 (first entry)

DE Human oxytocin secretion promoting peptide SEQ ID NO: 33.

XX Human; Oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.

OS Homo sapiens.

XX  
 PN WO200038704-A1.

XX  
 PD 06-JUL-2000.

XX  
 PE 22-DEC-1999; 99WO-JP07199.

XX  
 PR 25-DEC-1998; 98JP-0369585.

XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.

XX  
 PI Matsumoto H, Kitada C, Hinuma S;

XX  
 DR WPI; 2000-452298/39.

XX  
 PT Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine.

XX  
 PS Disclosure; Page 62; 72pp; Japanese.

XX  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC, which contains a ligand peptide or its salt for the G protein-coupled  
 receptor protein. It is useful in the form of drugs for ameliorating,

GenCore version 4.5  
 Copyright (c) 1993 - 2000 Compugen Ltd.

protein - protein search, using sw model

on:

April 17, 2001, 15:45:59 ; Search time 70.09 seconds  
 (without alignments)  
 21.574 Million celi updates/sec

query: US-09-446-543A-73

score: 109

oring table: BLASTp62

Gapop 10.0 , Gapext 0.5

searched: 198801 seqs, 68722935 residues

total number of hits satisfying chosen parameters: 198801

minimum DB seq length: 0

maximum DB seq length: 200000000

1st-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

database : PIR 67.0

1: Pirl:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

pred. No. 1s the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	49.5	45.4	664	F83376 conserved hypothetical protein PA2151. [imported] - <i>Pseudomonas aeruginosa</i> (strain PAO C;Species: <i>Pseudomonas aeruginosa</i> C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000 C;Accession: F83376 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L; 'Iony, S.; Olson, M.V.; Nature 406, 959-964, 2000 A;Title: Complete genome sequence of <i>Pseudomonas aeruginosa</i> PA01, an opportunistic pa A;Reference number: A82950; MUID:2043737 A;Accession: F83376 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-664 <ST0> A;Cross-references: GB:AE004642; GB:AE004091; NID:99948163; PIDN:AG05539.1; GSPDB:GN A;Experimental source: strain PA01 C;Genetics: A;Gene: PA2151
2	49.5	45.0	790	T47959 hypothetical protein hydroxybenzoate oc
3	46.2	333	2	H88852 hypothetical prote
4	41.3	276	T21969 hypothetical prote	
5	40.4	250	2	G83400 hypothetical prote
6	40.4	284	2	F71015 hypothetical prote
7	39.9	1501	2	T5623 hypothetical prote
8	43	394	220	C8292 probable glutathio
9	43	39.4	309	T2376 hypothetical prote
10	43	39.4	2	A48197 opsin, ocular - A
11	43	39.4	376	T47959 hypothetical protein PA2151. [imported] - <i>Pseudomonas aeruginosa</i> (strain PAO C;Species: <i>Pseudomonas aeruginosa</i> C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000 C;Accession: F83376 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L; 'Iony, S.; Olson, M.V.; Nature 406, 959-964, 2000 A;Title: Complete genome sequence of <i>Pseudomonas aeruginosa</i> PA01, an opportunistic pa A;Reference number: A82950; MUID:2043737 A;Accession: F83376 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-664 <ST0> A;Cross-references: GB:AE004642; GB:AE004091; NID:99948163; PIDN:AG05539.1; GSPDB:GN A;Experimental source: strain PA01 C;Genetics: A;Gene: PA2151
12	43	39.4	376	T48197 opsin, lateral eye
13	43	39.4	2	G64720 probable amino aci
14	43	39.4	719	S61046 ARPI protein - yea
15	43	39.4	1296	T16859 hypothetical prote
16	42.5	39.0	443	T21499 hypothetical prote
17	42	38.5	428	T8694 pyruvate dehydroge
18	41.5	38.1	345	B84012 N-acetylglutamate
19	41	37.6	2	S7955 hypothetical prote
20	41	37.6	264	A75117 hypothetical prote
21	41	37.6	338	T20100 hypothetical prote
22	41	37.6	342	B63395 malic acid transpo
23	41	37.6	343	T46534 probable FMN-depen
24	41	37.6	347	H63371 malt acid transpo
25	41	37.6	986	1 OYURGA spexact receptor p
26	41	37.6	2155	2 T30197 alpha tectorin -
27	40.5	37.2	779	2 T49117 related to BCS1 Pr
28	40.5	37.2	1540	2 T45619 hypothetical prote
29	40.5	36.7	184	2 T35041 probable membrane

RESULTS

Result No.	Query Match	Best Local Similarity	Score	DB	Length
1	45.4%	62.5%	49.5;	2;	664;
2	45.4%	62.5%	49.5;	2;	664;
3	45.4%	62.5%	49.5;	2;	664;
4	45.4%	62.5%	49.5;	2;	664;
5	45.4%	62.5%	49.5;	2;	664;
6	45.4%	62.5%	49.5;	2;	664;
7	45.4%	62.5%	49.5;	2;	664;
8	45.4%	62.5%	49.5;	2;	664;
9	45.4%	62.5%	49.5;	2;	664;
10	45.4%	62.5%	49.5;	2;	664;
11	45.4%	62.5%	49.5;	2;	664;
12	45.4%	62.5%	49.5;	2;	664;
13	45.4%	62.5%	49.5;	2;	664;
14	45.4%	62.5%	49.5;	2;	664;
15	45.4%	62.5%	49.5;	2;	664;
16	45.4%	62.5%	49.5;	2;	664;
17	45.4%	62.5%	49.5;	2;	664;
18	45.4%	62.5%	49.5;	2;	664;
19	45.4%	62.5%	49.5;	2;	664;
20	45.4%	62.5%	49.5;	2;	664;
21	45.4%	62.5%	49.5;	2;	664;
22	45.4%	62.5%	49.5;	2;	664;
23	45.4%	62.5%	49.5;	2;	664;
24	45.4%	62.5%	49.5;	2;	664;
25	45.4%	62.5%	49.5;	2;	664;
26	45.4%	62.5%	49.5;	2;	664;
27	45.4%	62.5%	49.5;	2;	664;
28	45.4%	62.5%	49.5;	2;	664;
29	45.4%	62.5%	49.5;	2;	664;

ALIGMENTS

Query Match 45.0%; Score 49; DB 2; Length 790;  
 Best Local Similarity 47.4%; Pred. No. 3.7; Mismatches 3; Indels 0; Gaps 0;  
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PDINPAWYXXKGIRPVGRF 20  
 Db 366 PPNNPRTYGSRQLOPHGRW 384

RESULT 3  
 H82852 hydroxybenzoate octaprenyltransferase xF0068 [Imported] - *Xylella fastidiosa* (strain 9as)  
 C;Species: *Xylella fastidiosa*  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
 C;Accession: H82852  
 R;anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen  
 Nature 406, 151-157, 2000  
 A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
 A;Reference number: A82515; MUID:20365717  
 A;Note: for a complete list of authors see reference number A59328 below  
 A;Accession: H82852  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-333 <SM>  
 A;Cross-references: GB:AE003860; GB:AE003849; NID:99104830; PIDN:AAF82881.1; GSDB:GN001  
 A;Experimental source: strain 9as  
 R;SIMSON, A.J.G.; Reinach, F.C.; Almeida, P.; Abreu, F.A.; Acencio, M.; Alvaranga, R.;  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrasco, D.M.; Carter, H.  
 ab-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to Genbank, June 2000  
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, M.C.; Frohn,  
 J.D.; Junqueira, M.L.; Kemper, S.L.; Kite, J.M.; Klieger, J.P.; Kurama, E.B.; Laij, G.;  
 Chado, M.A.; Madeira, A.M.B.N.; Madalena, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, H.  
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
 P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.;  
 Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki,  
 M.; Tschakos, M.H.; Valida, H.; van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;  
 M.; Reference number: A59328  
 A;Contents: annotation  
 C;Genetics  
 A;Gene: xF0068  
 C;Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 42.2%; Score 46; DB 2; Length 333;  
 Best Local Similarity 53.3%; Pred. No. 4.9; Mismatches 8; Indels 0; Gaps 0;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 INPAWYXXRG1RPVG 18  
 Db 54 LDPYWKLARGDRPVG 68

RESULT 4  
 T21969 hypothetical protein F38E11.7 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C;Accession: T21969  
 R;Matthews, P.  
 A;Accession: T21969  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-767 <WIL>  
 A;Cross-references: EMBL:268342; PIDN:CAA92775.1; GSPDB:GN00022; CESP:F38E11.7  
 A;Experimental source: clone F38E11  
 C;Genetics:  
 A;Gene: CESP:F38E11.7  
 A;MQP position: 4 118/1; 139/2; 189/3; 226/1; 248/1; 287/2; 375/2; 432/3; 465/3; 548/1; 4

Query Match 41.3%; Score 45; DB 2; Length 767;  
 Best Local Similarity 75.0%; Pred. No. 16; Mismatches 6; Conservative 1; Indels 1; Gaps 0;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 PDINPAWY 9  
 Db 747 PDVKPAWY 754

RESULT 5  
 G83400 hypothetical protein PA1952 [Imported] - *Pseudomonas aeruginosa* (strain PAO1)  
 C;Species: *Pseudomonas aeruginosa*  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C;Accession: G83400  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;  
 adman, S.; Yuan, Y.; Brody, L.I.; Coulter, S.N.; Folger, K.R.; Kas, A.; Iarbigi, K.; L.  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pa  
 A;Reference number: A82950; MUID:2043737  
 A;Accession: G83400  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-250 <SM>  
 A;Cross-references: GB:AE04622; GB:AE004091; NID:99947948; PIDN:ANG05340.1; GSDB:GN  
 A;Experimental source: strain PAO1  
 C;Genetics:  
 A;Gene: PA1952

Query Match 40.4%; Score 44; DB 2; Length 250;  
 Best Local Similarity 47.8%; Pred. No. 7.8; Mismatches 8; Indels 2; Gaps 0;  
 Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PDINPAWYXXKGIRPVGRF 18  
 Db 233 PALNCAWEQQLALRPG 249

RESULT 6  
 F71015 hypothetical protein PH1420 - *Pyrococcus horikoshii*  
 C;Species: *Pyrococcus horikoshii*  
 C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C;Accession: F71015  
 R;Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haliwawa, Y.; Hino, Y.; Yamamoto, S.; se  
 M.; Ohfuki, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Ogu  
 DNA Res. 5, 55-6, 1998  
 A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A;Reference number: A71000; MUID:98344137  
 A;Accession: F71015  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-284 <KAW>  
 A;Cross-references: GB:AP000006; NID:g3236133; PIDN:BAA30526.1; PID:g3257843  
 A;Experimental source: strain OT3  
 A;Note: this accession replaces an interim accession for a sequence replaced by GenBa  
 C;Genetics:  
 A;Gene: PH1420  
 C;superfamily: *Pyrococcus horikoshii* hypothetical protein PH1420

Query Match 40.4%; Score 44; DB 2; Length 284;  
 Best Local Similarity 44.4%; Pred. No. 8.9; Mismatches 8; Indels 0; Gaps 0;  
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 2 PDINPAWYXXKGIRPVGRF 19  
 Db 217 PYIEPPFYALRGLIELGR 224

RESULT 7

T45623 hypothetical protein F13G24.180 - *Arabidopsis thaliana* (mouse-ear cress)  
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
 C;Accession: T45623  
 R;Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; BE submitted to the Protein Sequence Database, December 1999  
 A;Reference number: Z23009  
 A;Accession: T45623  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-212  
 A;Cross-references: <NHLI>  
 A;Experimental source: strain R1  
 C;Genetics:  
 A;Map position: 5  
 A;Introns: 64/1; 739/3; 785/2; 1302/2; 1318/3; 1399/2; 1434/2  
 A;Note: F13G24.180

Query Match 39.9%; Score 43.5%; DB 2; Length 1501;  
 Best Local Similarity 34.5%; Pred. No. 57;  
 Matches 10; Conservative 5; Mismatches 5; Indels 9; Gaps 2;

Qy 1 TPDINPAWYXXRGIRPV--IRPV--GRF 20  
 Qy 2 PDINPAWYXXRGIRPV 16  
 Db 1184 SPQMAPSWISQYGTWKNGLVQPVNDTGRF 1212

RESULT 8

C83292 probable glutathione S-transferase PA2821 [Imported] - *Pseudomonas aeruginosa* (strain PA)  
 C;Species: *Pseudomonas aeruginosa*  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C;Accession: C83292  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Bradian, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lai, B.; Lim, J.; Lory, S.; Olson, M.V.  
 Nature 405, 955-964, 2000  
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen  
 A;Reference number: A82950; MUID:20437337  
 A;Accession: C83292  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-220 <NTO>  
 A;Experimental source: strain PA01  
 A;Cross-references: GB:AB004709; GB:AE004091; NID:99946904; PIDN:AA06209.1; GSDB:GN001  
 A;Genetics:  
 A;Gene: PA2821

Query Match 39.9%; Score 43.5%; DB 2; Length 1501;  
 Best Local Similarity 34.5%; Pred. No. 57;  
 Matches 10; Conservative 5; Mismatches 5; Indels 9; Gaps 2;

Qy 1 TPDINPAWYXXRGIRPV--IRPV--GRF 20  
 Qy 2 PDINPAWYXXRGIRPV 16  
 Db 1184 SPQMAPSWISQYGTWKNGLVQPVNDTGRF 1212

RESULT 9

G75608 hypothetical protein - *Deinococcus radiodurans* (strain R1)  
 C;Species: *Deinococcus radiodurans*  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C;Accession: G75608  
 R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Me S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 266, 1571-1577, 1999  
 A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

Query Match 39.4%; Score 43%; DB 2; Length 376;

RESULT 10

T32376 hypothetical protein K10F12.4 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C;Accession: T32376  
 R;Wohldmann, P.; Beck, C.  
 Submitted to the EMBL Data Library, September 1997  
 A;Description: The sequence of *C. elegans* cosmid K10F12.  
 A;Reference number: Z22157  
 A;Accession: T32376  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-309 <NCH>  
 A;Cross-references: EMBL:AE025462; PIDN:AB71002.1; GSDB:GN00021; CESP:K10F12.4  
 A;Experimental source: strain Bristol N2; clone K10F12  
 C;Genetics:  
 A;Gene: CESP:K10F12.4  
 A;Map position: 3  
 A;Introns: 31/3; 123/2; 196/3; 239/1

Query Match 39.4%; Score 43%; DB 2; Length 309;  
 Best Local Similarity 44.4%; Pred. No. 14;  
 Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

Qy 1 TPDINPAWYXXRGIRPV--IRPV--GRF 20  
 Qy 2 PDINPAWYXXRGIRPV 16  
 Db 131 PDRSPNWLPPS--PIGR 146

RESULT 11

A48197 opsins ocellar - Atlantic horseshoe crab  
 C;Species: *Limulus polyphemus* (Atlantic horseshoe crab)  
 C;Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 13-Aug-1999  
 C;Accession: A48197  
 R;Smith, W.C.; Price, D.A.; Greenberg, R.M.; Battelle, B.A.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 6150-6154, 1993  
 A;Title: Opsins from the lateral eyes and ocelli of the horseshoe crab, *Limulus polyphemus*  
 A;Reference number: A48197; MUID:93317641  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-316 <SNL>  
 A;Cross-references: EMBL:103192; NID:9156644; PIDN:AA28274.1; PID:9156645  
 C;Superfamily: vertebrate rhodopsin  
 C;Keywords: chromoprotein; G protein-coupled receptor; lipoprotein; photoreceptor; re F;318/Binding site: retinal (Iys) (covalent) #status predicted

Query Match 39.4%; Score 43%; DB 2; Length 376;

Best Local Similarity 46.2%; pred. No. 17;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0; Qy 3 DIN----PAWYXARGI 14  
Oy 2 PDIINAWIXXARGI 14  
Db 40 PPMNFIWYSILGV 52

RESULT 12  
B48197 opsin, lateral eye - Atlantic horseshoe crab  
C;Species: Limulus polyphemus (Atlantic horseshoe crab)  
C;Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 13-Aug-1999  
C;Accession: B48197  
R;Smith, W.C.; Price, D.A.; Greenberg, R.M.; Battelle, B.A.  
PROC. Natl. Acad. Sci. U.S.A. 90, 6150-6154, 1993  
A;Title: Opsins from the lateral eyes and ocelli of the horseshoe crab, *Limulus polyphemus*  
A;Reference number: A48197; MUID:93117611  
A;Accession: B48197  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-376 <ENI>  
A;Cross-references: EMBL:103791; NID:9156642; PIDN:AAA28273.1; PID:9156643  
C;Superfamily: vertebrate rhodopsin  
C;Keywords: Chromoprotein; G protein-coupled receptor; lipoprotein; photoreceptor; retinol-binding site; retinal (Lys) (covalent) #status predicted

Query Match 39.4%; Score 43; DB 2; Length 376;  
Best Local Similarity 46.2%; pred. No. 17;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0; Qy 2 PDINAWYXARGI 14  
Db 40 PPMNFIWYSILGV 52

RESULT 13  
G64720 probable amino acid transport protein *yaaj*, sodium-dependent - *Escherichia coli*  
C;Species: *Escherichia coli*  
C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 26-Aug-1999  
C;Accession: G64720  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; McClelland, M.; Rose, D.J.; Mau, B.; Shao, Y.; Science, 277, 153-162, 1997  
A;Title: The complete genome sequence of *Escherichia coli* K-12.  
A;Reference number: A64720; MUID:91426617  
A;Accession: G64720  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-176 <BLAT>  
A;Cross-references: EMBL:274215; NID:91431265; PIDN:CAA98741.1; PID:e253076; PID:91431265  
A;Experimental source: strain S288C  
C;Genetics: SGD:NRPL; ARPL  
A;Cross-references: MIPS:YDL167c; SGD:S002326  
A;Map position: 4L

Query Match 39.4%; Score 43; DB 2; Length 719;  
Best Local Similarity 50.0%; pred. No. 33;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0; Qy 7 AWYXXRGIRDVG 18  
Db 244 SWFTQYGVRPVG 255

RESULT 15  
T116859 hypothetical protein *T11c2.5* - *Caenorhabditis elegans*  
C;Species: *Caenorhabditis elegans*  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Nov-2000  
C;Accession: T116859  
R;Du, Z.; submitted to the EMBL Data Library, November 1995  
A;Description: The sequence of *C. elegans* cosmid T11c2.  
A;Reference number: Z118591  
A;Accession: T116859  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1296 <DUZ>  
A;Cross-references: EMBL:U40030; NID:9105164; PIDN:AAA81133.1; CESP:T116859  
C;Genetics: CESP:T11c2.5  
A;Introns: 62/3; 93/3; 124/3; 142/1; 212/1; 293/1; 336/2; 406/3; 437/1; 508/1; 584/1;  
Query Match 39.4%; Score 43; DB 2; Length 1296;

Best Local Similarity 45.0%, pred. No. 60;  
Matches 9; Conservative 2; Mismatches 7;  
Indels 2; Gaps 1;  
Qy 1 TPPINPAWYXARGIRGVRF 20  
| || | || : : |  
Db 975 TPPINSDWYFSR--ADINRF 992

search completed: April 17, 2001, 15:45:59  
Job time: 607 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM Protein - Protein search, using sw model

Run on: April 17, 2001, 15:48:53 ; Search time 39.1 Seconds  
(without alignments)  
19.274 Million cell updates/sec

Title: US-09-446-543a-73

Perfect score: 109  
Sequence: 1 TPDINPAWYXRG1RPVGRFXX 22

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	105	96.3	87	1	PRRP_HUMAN	P81277	homo sapien
2	104	95.4	83	1	PRRP_BOVIN	P81278	rattus norvegicus
3	104	95.4	98	1	PRRP_RAT	P81264	bos taurus
4	46	42.6	676	1	EXKL_HUMAN	P92935	homo sapien
5	43	42.4	376	1	OSRL_LIMPO	P35360	limulus polyphemus
6	43	39.4	376	1	CICR_CHRVI	P35361	limulus polyphemus
7	43	39.4	383	1	Y493_WYCTU	P02947	chromatium
8	43	39.4	476	1	VAJ_ECOLI	P30143	escherichia coli
9	43	39.4	719	1	ABP_YEAST	P32770	saccharomyces cerevisiae
10	41	37.6	342	1	Y762_MEJVA	Q58172	methanococcus jannaschii
11	41	37.6	347	1	Y756_MEJVA	Q57995	methanococcus jannaschii
12	41	37.6	986	1	CIGR_ARBU	P11528	arbacia punctulata
13	40	36.7	329	1	Y493_WYCTU	Q11538	mycobacterium
14	40	36.7	546	1	CHOD_SR5Q	P12675	streptomyces
15	40	36.7	581	1	POE_MJYK	P31795	radiation
16	40	36.7	843	1	POE_MJYAK	P03357	murine
17	40	36.7	1196	1	POE_MJYAV	P03355	akr murine
18	40	36.7	1196	1	POE_MJYRD	P11227	radiation
19	39.5	36.2	860	1	VS12_BPEB03	P07893	bacteriophaga
20	39	35.8	149	1	ERBM_BEP7	P09541	bacteriophaga
21	39	35.8	360	1	WWR2_HUMAN	P09543	homo sapien
22	39	35.8	377	1	OPSL_HENSA	Q25157	hemigrapsus
23	39	35.8	377	1	OPSL_HENSA	Q25158	hemigrapsus
24	39	35.8	622	1	PECC_RAT	P07379	rattus norvegicus
25	39	35.8	953	1	SYV_VIRCH	Q9kp73	vibrio cholerae
26	39	35.8	952	1	UBRA_MENTH	Q26543	methanobacter
27	38.5	35.3	1	PLSC_HELPY	Q25903	helicobacte	
28	38.5	35.3	265	1	UBIE_RICPR	Q92cp3	rickettsia
29	38.5	35.3	831	1	DPOL_THEFL	P30113	thermus aquitum
30	38.5	35.3	877	1	CAD2_BOVIN	P19534	bos taurus
31	38.5	35.3	906	1	CAD2_HUMAN	P19022	homo sapien
32	38.5	35.3	906	1	CAD2_MOUSE	P15116	mus musculus
33	38.5	35.3	3083	2	POLG_2HMVR	Q89330	genome po

## ALIGNMENTS

34	38	34.9	158	1	RJ15_AERPE	Q9Y98 aeropyrum
35	38	34.9	197	1	T1BB_HUMAN	Q95998 homo sapien
35	38	34.9	261	1	PHSC_ECOLI	P77409 escherichia
37	38	34.9	332	1	LITB_MICCLE	Q9x781 mycobacteri
38	38	34.9	399	1	HMPA_BACSU	P49852 bacillus su
39	38	34.9	951	1	SYV_ECOLI	P07118 escherichia
40	38	34.9	994	1	GNE_MCTU	Q10379 mycobacteri
41	38	34.9	1723	1	ATML_HUMAN	Q9y4k1 homo sapien
42	38	34.9	2870	1	M0VE_RHIME	P20471 rhizobium m
43	37	33.9	179	1	RK6_GUTTH	Q46908 guillardia
44	37	33.9	251	1	VP24_EBOM4	Q11459 ebola virus
45	37	33.9	251	1	VP24_EBOM2M	Q05322 ebola virus





RL PROC: NATL Acad. Sci. U.S.A. 90:6150-6154 (1993).  
 CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT  
 CC MEDIATE VISION. THEY CONSIST OF AN APPROPRIATE OPSIN, COVALENTLY  
 CC LINKED TO CIS-RETINYL.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: OCULAR CELLS; MEDIAN OCULI.  
 CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY  
 CC BE PHOSPHORYLATED (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMA AT 530 NM.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC OPSIN SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-ilb.ch/announce/> or send an email to license@isb-ilb.ch).

CC EMBL: L03792; AAA2B274.I; -.  
 CC PIR: A48197; AA8197.  
 DR GCRDB; GCR\_05B5; -.  
 DR InterPro: IPR002776; -.  
 DR InterPro: IPR001391; -.  
 DR InterPro: IPR001760; -.  
 DR PFAM: PF00001; 7tm\_1; 1.  
 DR PRINTS: PR00237; GCPFRHOPPSN.  
 DR PRINTS: PR00238; OPSIN.  
 DR PRINTS: PR00578; OPSINTRUEYE.  
 DR PROSITE: PS50262; G-PROTEIN\_RCCEP\_FL1; I.  
 DR PROSITE: PS00238; OPSIN; I.  
 KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision; phosphorylation; G-protein coupled receptor.

CC PT DOMAIN 1 46 EXTRACELLULAR.  
 CC PT TRANSMEM 47 71 (POTENTIAL).  
 CC PT DOMAIN 72 83 CYTOSMERIC.  
 CC PT TRANSMEM 84 108 2 (POTENTIAL).  
 CC PT DOMAIN 109 123 EXTRACELLULAR.  
 CC PT TRANSMEM 124 143 3 (POTENTIAL).  
 CC PT DOMAIN 144 162 CYTOPLASMIC.  
 CC PT TRANSMEM 163 186 4 (POTENTIAL).  
 CC PT DOMAIN 187 210 EXTRACELLULAR.  
 CC PT TRANSMEM 211 238 5 (POTENTIAL).  
 CC PT DOMAIN 239 274 CYTOPLASMIC.  
 CC PT TRANSMEM 275 298 6 (POTENTIAL).  
 CC PT DOMAIN 299 306 EXTRACELLULAR.  
 CC PT TRANSMEM 307 331 7 (POTENTIAL).  
 CC PT DOMAIN 332 376 CYTOPLASMIC.  
 CC PT DISULFID 120 197 BY SIMILARITY.  
 CC PT BINDING 318 318 RETINAL CHROMOPHORE (BY SIMILARITY).  
 CC PT CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC PT CARBOHYD 193 193 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC SEQUENCE 376 AA; 42111 MW; FA9647C4053TCBP8 CRC64;

Query Match 39.4%; Score 43; DB 1; Length 376;  
 Best Local Similarity 46.2%; Pred. No. 8.5; Mismatches 5; Indels 0; Gaps 0;  
 Matches 5; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

RESULT 7  
 CYCR\_CHRVT STANDARD; PRT; 383 AA.  
 QY 2 PDIHAWYXRG1 14  
 ID 1 :||| 1 :| :|  
 DB 40 PPMHPLWYSTIGV 52

DT 30-MAY-2000 (REL. 39, Last annotation update)  
 DE PHOTOSYNTHETIC REACTION CENTER CYTOCHROME C SUBUNIT PRECURSOR.  
 DE PUF.C.  
 OS Chromatium vinosum.  
 OC Bacteria; Prokaryotes; gamma subdivision; Chromatiales; Allocyano; OC  
 NCBI TAXID=1049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Corson G.P., Nagashima K.Y., Matsuura K., Sakuragi Y., Ruwanthi W.,  
 RA Qin H., Allen R., Knaff D.B.;  
 RA "Primary structure of genes encoding light-harvesting and reaction center proteins from Chromatium vinosum.";  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases.  
 CC -1- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO OXIDIZED PRIMARY ELECTRON DONOR.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID ANCHOR (BY SIMILARITY).  
 CC -1- PTM: BINDS FOUR HEME GROUPS PER MOLECULE.  
 CC -1- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER CYTOCHROME C SUBUNITS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-ilb.ch/announce/> or send an email to license@isb-ilb.ch).

CC EMBL: AB01181; BAA32742.I; -.  
 DR PROSITE; PS01909; CYTOCHROME\_C; 4.  
 DR InterPro: IPR003045; -.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 4.  
 KW Membrane; Lipoprotein; Duplication; Signal; SIGNAL\_I; 22 BY SIMILARITY.  
 CC PT CHAIN 23 383 PHOTOSYNTHETIC REACTION CENTER CYTOCHROME C SUBUNIT.  
 CC PT LIPID 23 23 N-ACYL DIGLYCERIDE (BY SIMILARITY).  
 CC PT BINDING 107 107 HEME 1 (COVALENT) (BY SIMILARITY).  
 CC PT BINDING 110 110 HEME 1 (COVALENT) (BY SIMILARITY).  
 CC PT METAL 111 111 IRON 1 (HEME AXIAL LIGAND)  
 CC PT METAL 152 152 HEME 2 (COVALENT) (BY SIMILARITY).  
 CC PT BINDING 155 155 HEME 2 (COVALENT) (BY SIMILARITY).  
 CC PT METAL 156 156 IRON 2 (HEME AXIAL LIGAND)  
 CC PT BINDING 247 247 HEME 3 (COVALENT) (BY SIMILARITY).  
 CC PT BINDING 250 250 HEME 3 (COVALENT) (BY SIMILARITY).  
 CC PT METAL 251 251 IRON 3 (HEME AXIAL LIGAND)  
 CC PT BINDING 307 307 HEME 4 (COVALENT) (BY SIMILARITY).  
 CC PT BINDING 310 310 HEME 4 (COVALENT) (BY SIMILARITY).  
 CC PT METAL 311 311 IRON 4 (HEME AXIAL LIGAND)  
 CC SEQUENCE 383 AA; 41522 MW; 96BCD918FB9A87E CRC64;

Query Match 39.4%; Score 43; DB 1; Length 383;  
 Best Local Similarity 36.8%; Pred. No. 8.6; Mismatches 9; Indels 0; Gaps 0;  
 Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

RESULT 8  
 YAAJ\_ECOLI STANDARD; PRT; 476 AA.  
 QY 1 TPDINPAWYXRG1PVRG 19  
 ID YAAJ\_ECOLI  
 DB 263 TPDINPAWYXRG1PVRG 281  
 AC P30143;





CC -I- CATALYTIC ACTIVITY: GTP = 3',5'-CYCLIC GMP + PYROPHOSPHATE.  
 CC -I- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -I- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC  
 CC DOMAIN OF PROTEIN KINASES.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X12874; CAA31367.1; -.

DR PIR: S05480; OXURGA; -.

DR InterPro: IPR00719; -.

DR InterPro: IPR01628; -.

DR Pfam: PF0094; ANP; receptor; 1.

DR Pfam: PF00669; Pkinase; 2.

DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.

KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase; CGMP synthesis; Signal.

FT SIGNAL 1 21

FT CHAIN 22

FT DOMAIN 22 507 REACT RECEPTOR.

FT TRANSM 508 528 EXTRACELLULAR (POTENTIAL).

FT DOMAIN 529 986 CYTOSOLIC (POTENTIAL).

FT DOMAIN 568 836 PROTEIN KINASE LIKE.

FT CARBOND 185 185 N-LINKED (GlcNAc. . .) (POTENTIAL).

FT CARBOND 361 361 N-LINKED (GlcNAc. . .) (POTENTIAL).

FT CARBOND 410 410 N-LINKED (GlcNAc. . .) (POTENTIAL).

SO SEQUENCE 986 AA; 111284 MW; B40238A74CCAF52 CRC64;

Query Match 37.6%; Score 41; DB 1; Length 986; Best Local Similarity 62.5%; Pred. No. 46; Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0; Oy 2 PDINPAWY 9 Db 475 PDINPAWYH 482

RESULT 13

Y493\_MCTU STANDARD; PRT: 329 AA.

AC 0115; DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE HYPOTHETICAL PROTEIN RV0493C.

GN RV0493C OR MTCY20G9-19C.

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinomycetales; Streptomyces; Actinomycetales; Streptomyceae; Streptomyces; Streptomyces sp. (strain SA-COO).

OC NCBI\_TAXID-1931; OX NCBI\_TAXID-1931; RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINV-89123081; Pubmed-2914858;

RX MEDLINV-89211873; Pubmed-10194345;

RA Ishizaki T., Hirayama N., Shinkawa H., Nimi O., Murooka Y.;  
 "Nucleotide sequence of the gene for cholesterol oxidase from a  
 RT Streptomyces sp.";

RT J. Bacteriol. 171:596-601(1989).

[2]

RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS), AND MUTAGENESIS.

RX MEDLINV-99211873; Pubmed-10194345;

RA Yue O.K., Nass I.J., Sampson N.S., Vrielink A.;

RT "Crystal structure determination of cholesterol oxidase from Streptomyces and structural characterization of key active site mutants";

RT Biochemistry 38:4277-4286(1999).

CC -I- CATALYTIC ACTIVITY: CHOLESTEROL + O(2) = CHOLEST-4-EN-3-ONE + H<sub>2</sub>O(2).

CC -I- COFACTOR: FAD.

CC -I- PATHWAY: CHOLESTEROL METABOLISM.

CC -I- SUBUNIT: MONOMER.

CC -I- SIMILARITY: BELONGS TO THE GMC\_OXIDOREDUCTASES FAMILY.

CC STRAIN-TAXID; RN SEQUENCE FROM N.A.

RX MEDLINV-98295987; Pubmed-9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Elbrügger K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrell B.G., Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature 393:537-544(1998).

-I- SIMILARITY: SOME, TO MLEPRAB B2168\_F2\_93.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: Z77162; CAA00954.1; -.

DR DR Tuberculist; RV0493C; -.

KW Hypothetical protein.

SO SEQUENCE 329 AA; 35427 MW; BE4B524750277B56 CRC64;

Query Match 36.7%; Score 40; DB 1; Length 329; Best Local Similarity 42.9%; Pred. No. 23; Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0; Oy 2 PDINPAWYXKGIR 15 Db 90 RAAGPAWFDIAGVR 103

RX CHOD\_SRSQ STANDARD; PRT: 546 AA.

AC P12676; DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DR CHOLESTEROL OXIDASE PRECURSOR (EC 1.1.3.6) (CHOD).

GN CHOA.

OS Streptomyces sp. (strain SA-COO).

OC Bacteria; Firmicutes; Actinobacteria; Actinomycetales; Streptomyces; Actinomycetales; Streptomyceae; Streptomyces.

OC NCBI\_TAXID-1931; OX NCBI\_TAXID-1931; RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINV-89123081; Pubmed-2914858;

RX MEDLINV-99211873; Pubmed-10194345;

RA Ishizaki T., Hirayama N., Shinkawa H., Nimi O., Murooka Y.;  
 "Nucleotide sequence of the gene for cholesterol oxidase from a  
 RT Streptomyces sp.";

RT J. Bacteriol. 171:596-601(1989).

[2]

RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS), AND MUTAGENESIS.

RX MEDLINV-99211873; Pubmed-10194345;

RA Yue O.K., Nass I.J., Sampson N.S., Vrielink A.;

RT "Crystal structure determination of cholesterol oxidase from Streptomyces and structural characterization of key active site mutants";

RT Biochemistry 38:4277-4286(1999).

CC -I- CATALYTIC ACTIVITY: CHOLESTEROL + O(2) = CHOLEST-4-EN-3-ONE + H<sub>2</sub>O(2).

CC -I- COFACTOR: FAD.

CC -I- PATHWAY: CHOLESTEROL METABOLISM.

CC -I- SUBUNIT: MONOMER.

CC -I- SIMILARITY: BELONGS TO THE GMC\_OXIDOREDUCTASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: M31939; ALA26719.1; -.

DR PIR: A32260; A32260.

DR PDB: 1B4V; 06-JAN-99.

DR PDB: 1B8S; 09-FEB-99.

DR PDB: 1CBO; 10-MAR-99.

DR PDB: 1CC2; 11-MAR-99.

Query Match 36.7%; Score 40; DB 1; Length 546;  
 Best Local Similarity 36.8%; PScore, No. 38;  
 Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;  
 Qy 2 PDPINPAWXXRGIRPVGRF 20  
 Db 98 PDKRSSWPKNRTEAPLGSE 116

Search completed: April 17, 2001, 15:48:53  
 Job time: 541 sec

FT	NON_TER	1	1
SQ	SEQUENCE	581 AA:	65157 MW: 8D7A38694C8E036E CRC64;
Query Match	36.7%;	Score 40;	DB 1; Length 546;
Best Local Similarity	36.8%;	PScore, No. 38;	Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
Qy	1 PDPINPAWXXRGIRPVGRF 20	Db	556 TPPIRPSWRVQSNPL 572

RESULT 15  
 POL\_MVRK STANDARD; PRT; 581 AA.  
 ID POL\_MVRK  
 AC P31795;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE POLYPROTEIN [CONTAINS: PROTEASE (EC 3.4.23.-); REVERSE  
 DE TRANSCRIPTASE (EC 2.7.7.49); RIBONUCLEASE H (EC 3.1.26.4)] (FRAGMENT).  
 GN POL.  
 OS Radiation murine leukemia virus (strain Kaplan).  
 OC Viruses; Retrovirus; Retroviridae; Mammalian type C retroviruses.  
 OX NCBI\_TaxID:31689;  
 RN [1]  
 RP SEQUENCE FROM R. A.  
 RX MEDLINE=92333703; PUBMED=1629959;  
 RA Rassart E.;  
 RT Poliquin L., Bergeron D., Fortier J. L., Paquette Y., Bergeron R.,  
 RT "Determinants of thymotropism in Kaplan radiation leukemia virus and  
 RT nucleotide sequence of its envelope region.";  
 RL J. Virol. 66:511-5146 (1992).  
 CC -I- PRM: SPECIFIC ENZYMIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -I- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-POL  
 CC POLYPROTEIN.  
 CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE  
 CC RETROPEPSIN FAMILY.  
 CC  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is not modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC  
 EMBL; M93052; AAA46525.1; -.  
 PIR; A42743; A42743.  
 MEROPS; A02; 008; -.  
 InterPro; IPR001037; -.  
 InterPro; IPR001584; -.  
 InterPro; IPR01969; -.  
 InterPro; IPR00156; -.  
 DR pfam; PF00552; integrase; 1.  
 DR pfam; PF00073; nuclease; 1.  
 DR pfam; PF00065; rve; 1.  
 DR PROST; PS00141; ASP\_PROTEASE; PARTIAL.  
 KW Hydrolase; Transferase; RNA-directed DNA polymerase;  
 KW Aspartyl protease; Endonuclease; Polyprotein.



THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

On protein - protein search, using sw model

Run on: April 17, 2001, 15:48:13 ; Search time 115.07 Seconds

(without alignments)  
22.409 Million cell updates/sec

Title: US-09-446-543a-73

Perfect score: 109

Sequence: 1 TPDINPAWYXKRGIRPVGREGXX 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPREMBEL-15:  
1: sp\_archea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phage:\*

10: sp\_plant:\*

11: sp\_ridom:\*

12: sp\_unclassified:\*

13: sp\_vertebrate:\*

14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	82	75.2	117	13 Q9W624
2	49	45.0	790	Q9m371 arabidopsis
3	46	42.2	54	Q9UJF9 homo sapien
4	46	42.2	2	Q9PH76 xylella fastidiosa
5	46	42.2	4	Q9D677 homo sapien
6	46	42.2	10	Q9LG20 oryza sativa
7	45	41.3	767	5 020170
8	44	40.4	284	1 050128
9	44	40.4	419	4 Q9Y276
10	43.9	39.9	1501	10 Q9SD86
11	43	39.4	232	2 Q9RZ21
12	43	39.4	309	5 017234
13	43	39.4	1296	5 Q22452
14	42.5	39.0	333	2 Q9RJ0
15	42.5	39.0	443	5 Q19879
16	42	38.5	293	2 Q9PS55
17	42	38.5	428	2 Q9PK7
18	41.5	38.1	345	2 Q9K8V2
19	37.6	2	97	2 Q33440

ALIGNMENTS									
RESULT	1	Q9W624	Q9W624	PRELIMINARY;	PRT;	117 AA.			
ID	Q9W624	Q9W624	Q9W624;	1999 (TREMBEL. 12, Created)					
AC	Q9W624;	01-NOW-1999 (TREMBEL. 12, last sequence update)	DT	01-NOV-1999 (TREMBEL. 12, last annotation update)					
DT	01-NOW-1999 (TREMBEL. 12, last annotation update)		DT	01-NOW-1999 (TREMBEL. 12, last annotation update)					
DE			DE	C-RF AMIDE PRECURSOR.					
OS			OS	Carassius auratus (Goldfish).					
OC			OC	Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi; Actinopterygii; Neopteridae; Cyprinidae; Cyprininae; Carassius.					
OX			OX	Cypriniformes; Cyprinidae; Cyprininae; Carassius.					
RN			RN	NCBI TAXID=7957;					
RP			RP	[1]					
RC			RC	SEQUENCE FROM N.A.					
RA			RA	TISSUE-BRAIN;					
RT			RT	"Carassius" RFamide (C-RF amide).",					
RL			RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.					
EMBL;			EMBL;	AB020024; BAA76622; 1;					
SQ			SEQUENCE	117 AA; 12879 MW; D5DC4CB2203BC2B0 CRC64;					
Query Match 75.28; Score 82; DB 13; Length 117; Best Local Similarity 60.0%; Pred. No. 3.3e-06; Matches 12; Conservative 5; Mismatches 3; Indels 0; Gaps 0;									
QY	1	TPDINPAWYXKRGIRPVGREGXX	Db	56 SPEIDPFWVGRGVPIGRF	75				
RESULT	2	Q9M371	Q9M371	PRELIMINARY;	PRT;	790 AA.			
ID	Q9M371	Q9M371	Q9M371;	01-OCT-2000 (TREMBEL. 15, Created)	AC				
DT	01-OCT-2000 (TREMBEL. 15, last sequence update)		DT	01-OCT-2000 (TREMBEL. 15, last annotation update)	DT				
DE	HYPOTHETICAL	DE	DE	HYPOTHETICAL	DE				
GN	67.4 KDA PROTEIN.	GN	GN	F15G16.60.	GN				
OS	Arabidopsis thaliana (Mouse-ear cress).	OS	OS	Pseudomonas aeruginosa	OS				

OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Rosids; Brassicales; Brassicaceae; Arabidopsis.
OX	NCBI_TAXID=3702;
[1]	SEQUENCE FROM N.A.
RA	De Haan M., Marse A.C., Grivell L.A., Mewes H.W., Lemcke K., Meyer K.F.X., Querier F., Salanoubat M.; Submitted (Nov-1999) to the EMBL/GenBank/DBJ databases.
RP	[12] I-21.
RA	SEQUENCE FROM N.A.
RA	EU Arabidopsis sequencing project:
RA	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AL132954; CNB71071.1; -.
RA	Hypothetical protein:
RA	SEQUENCE 790 AA; 87375 MW; B22274B75690F30 CRC64;
RA	Query Match Score 49; DB 10; Length 790; Best Local Similarity 47.4%; Pred. No. 7.4; Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
Qy	2 PDINPAWXXRGRPVGR 20
Db	366 PPNPRTGSGRGLQPHGAW 384
RESULT 3	
Q9UJF9	PRELIMINARY; PRT; 54 AA.
AC	Q9UJF9: 01-MAY-2000 (TREMBl. 13, Created)
DT	01-MAY-2000 (TREMBl. 13, Last sequence update)
DR	01-MAY-2000 (TREMBl. 13, Last annotation update)
DE	DI479J7.3 (SUSHI-REPEAT PROTEIN (SRPUL)) (FRAGMENT).
GN	DI479J7.3.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCHI_TAXID=3606;
[1]	SEQUENCE FROM N.A.
RA	Lawlor S.; Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RL	NON-TER 54
DR	EMBL; AL035608; CAB55602.1; -.
FT	SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;
SQ	
RESULT 4	
Q9PH76	PRELIMINARY; PRT; 333 AA.
AC	Q9PH76; 01-OCT-2000 (TREMBl. 15, Created)
DT	01-OCT-2000 (TREMBl. 15, Last sequence update)
DE	DE HYDROXYBENZOATE OCTAPRENYLTRANSFERASE.
GN	XF0058
OS	Xylella fastidiosa.
OC	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group; Xylella.
OX	NCBI_TAXID=2371;
[1]	SEQUENCE FROM N.A.
RA	Query Match Score 42.2%; DB 4; Length 54; Best Local Similarity 43.8%; Pred. No. 1.3; Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;
Qy	1 TDINPAWXXRGRIP 16
Db	18 TRAVTPWNRAGSCTYV 33
RESULT 5	
Q9PH76	PRELIMINARY; PRT; 465 AA.
AC	Q9PH76; 01-AUG-1998 (TREMBl. 07, Created)
DT	01-AUG-1998 (TREMBl. 07, Last sequence update)
DT	01-MAY-2000 (TREMBl. 13, Last annotation update)
DE	SUSHI-REPEAT PROTEIN.
GN	SRPUL.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCHI_TAXID=9606;
RN	SEQUENCE FROM N.A.
RA	Kurokawa H., Inukai T., Inaba T., Goi K., Chang K.-S., Sinjyo T., Rakeshraw K.M., Naeve C.W., Look T.A.; Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RL	EMBL; AF00567; AAC15765.1; -.
DR	INTERPRO; IPR00436; -.
DR	INTERPRO; IPR00128; -.
DR	PFAM; PF00084; sushi; 3.
SO	SEQUENCE 465 AA; 52971 MW; 4D752B187FF3EB8 CRC64;
RN	[1]







DE F28D1.8 PROTEIN.  
 GN F28D1.8.  
 OC Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoldea;  
 OC Rhabditidae; Palaeoheriniae; Caenorhabditis;  
 OC NCBI\_TAXID=6239;  
 RN [1]  
 RP Baynes C.;  
 RA Baynes C.;  
 RL Submitted (APR-1995) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91150718; PubMed=7906398;  
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copepe T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Montimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Roopra K., Saunders D., Showkeen R.,  
 RA Smalton N., Smith A., Sonnhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkison-Spratt J., Wohldman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38 (1994).  
 DR EMBL: Z70084; CAA94603.1; -.  
 SQ SEQUENCE 443 AA; 48865 MW; D9137BAACCEF575E CRC64;

Query Match 39.0%; Score 42.5; DB 5; Length 443;  
 Best Local Similarity 37.5%; Pred. No. 47;  
 Matchers 9; Conservative 3; Mismatches 3; Indels 9; Gaps 1;

Qy 4 INP<sup>W</sup>YXXR-----GIRPVG 18  
 DB 53 INP<sup>W</sup>FDW<sup>W</sup>RVSVS<sup>W</sup>DIGHLG<sup>W</sup>PIG 76

Search completed: April 17, 2001, 15:48:13  
 Job time: 566 sec